

Создание общедоступных библиографических и полнотекстовых ресурсов по математике библиотеками РАН

Госина Людмила Игоревна, д-р филол. наук, ведущий научный сотрудник, Библиотека по естественным наукам Российской академии наук, филиал в Математическом институте им. В.А. Стеклова Российской академии наук, library@mi.ras.ru

Масляк Татьяна Ивановна, старший научный сотрудник, Библиотека по естественным наукам Российской академии наук, филиал в Математическом институте им. В.А. Стеклова Российской академии наук

Погорелко Константин Павлович, канд. техн. наук, ведущий научный сотрудник, Библиотека по естественным наукам Российской академии наук, филиал в Математическом институте им. В.А. Стеклова Российской академии наук

В докладе рассказывается об академических электронных ресурсах и корпоративных проектах, предусматривающих свободный доступ к библиографическим и полнотекстовым базам данных. Эти проекты направлены на создание универсальных по тематике или отраслевых (например, математических) библиографических и фактографических баз данных, полнотекстовых электронных библиотек, доступных всем пользователям в Интернете. Создателем математических ресурсов является содружество библиотек математических институтов РАН во главе с библиотекой Математического института им. В.А. Стеклова, в создании универсальных по тематике общедоступных ресурсов ведущая роль принадлежит БЕН РАН, ИНИОН, Архиву и другим академическим учреждениям.

В структуре ЦБС БЕН РАН уже более десяти лет реализуется проект создания отраслевой информационной системы по математике. Библиотечные сети РАН включают значительное число библиотек, объединенных в региональные централизованные сети. Глобальной централизации всех библиотек РАН никогда не существовало в силу их территориальной разделенности, сложностей с транспортом, различными каналами финансирования. Но работа библиотек координируется Информационно-библиотечным советом РАН под руководством одного из вице-президентов, и различные ЦБС сотрудничают в формировании и использовании фондов, создании электронных библиотек, научно-исследовательской работе. В современных условиях новые технологии позволяют организовать корпоративное взаимодействие библиотек из разных ЦБС и регионов по отраслевому принципу (по отрасли знания или комплексной проблеме). Первыми в этом направлении начали объединяться математические

библиотеки, вслед за ними приступили к созданию отраслевой электронной библиотеки физики. основополагающим принципом в создании этих систем является максимальная доступность информации пользователям в рамках действующего законодательства.

Библиотека Математического института им. В.А. Стеклова РАН, которую мы представляем, в сотрудничестве с академическими математическими библиотеками в других регионах формирует массив данных о своих фондах отечественных и зарубежных журналов и отечественных журнальных статьях по физико-математическим наукам. Основные отечественные профильные журналы распределены между четырьмя академическими партнерами, из которых 2 библиотеки (Математического института им. В.А. Стеклова РАН и Вычислительного центра им. А.А. Дородницына РАН) являются филиалами БЕН и находятся в Москве. Библиотека С.-Петербургского отделения Математического института им. В.А. Стеклова входит в систему БАН, а библиотека Института математики им. С.Л. Соболева находится в академгородке под Новосибирском и относится к централизованной сети ГПНТБ СО РАН. С нами начала работать и библиотека Института математики и механики Уральского отделения РАН (г. Екатеринбург), но с 2012 г. они лишились финансовой поддержки и их фонды пока не раскрываются в сводном каталоге.

Этот проект задумывался на рубеже 2000-х гг. и основывался на идее создания авторитетных, и в то же время общедоступных отраслевых информационных систем в едином информационном пространстве науки и образования. Информация об отечественных изданиях по математике интересна не только ученым РАН и профессорско-преподавательскому составу университетов и технических вузов, но также и учащейся молодежи, предполагающей продолжать обучение в аспирантуре и докторантуре. Корпус библиографических данных должен был распределиться между сводным каталогом журналов о фондах всех библиотек-участниц, сводным каталогом книг и базой данных «Статьи в отечественных математических журналах», охватывающей оглавления основных научных и научно-методических журналов. Договариваясь о сотрудничестве, мы не знали количественных показателей дублетности наших фондов, количества лакун и только предполагали, что спрос на эту информацию будет высоким.

В настоящее время сведения о фондах 4-х библиотек в сводный журнальный каталог уже практически полностью введены (исключение – небольшой объем журналов в Институте математики им. С.Л. Соболева, которая позже других библиотек вошла в проект). В настоящее время по каталогу числится 1527 названий журналов (100%), в том числе:

- иностранных - 1290 названий (84,5 %);
- отечественных – 237 названий (15,5 %).

По сводному каталогу за весь период с 19 века по настоящее время из 1527 названий лишь 242 издания (15,8 %) хранятся в одной из математических библиотек, в том числе 189 иностранных (12,4 %) и 53 отечественных (3,5 %). Цифру в 242 уникальных названия журналов примем за 100 % и посмотрим, как они распределены по фондам библиотек:

- 94 названия (39%) хранятся в библиотеке Математического института им. В.А. Стеклова (Москва);
- 87 названий (36%) хранятся в Институте математики им. С.Л. Соболева (Новосибирск).

Соответственно 31 математический журнал и 30 изданий по вычислительной математике хранятся в фондах библиотеки в С.-Петербурге библиотеки ВЦ в Москве.

В настоящее время сводный каталог математической литературы включает 637 названий журналов (100 %), поступивших в фонды библиотек-участниц за 2011-12 гг., из них 411 зарубежных (64,5 %) и 226 отечественных (35,5 %) изданий.

Дублетные издания последних двух лет (2011-2012 гг.) распределяются между библиотеками следующим образом:

- у всех 4-х библиотек из новых поступлений (в названиях) общими являются лишь 22 издания (4 иностранных и 18 отечественных),
- в трех библиотеках есть 52 общих журнала (33 иностранных и 19 отечественных),
- в двух библиотеках общими являются 78 изданий (56 иностранных и 22 отечественных),
- лишь в какой-то одной из 4-х библиотек имеются в наличии 237 названий (184 иностранных и 53 отечественных) и эти единственные экземпляры составляют 37 % от поступлений 2011-2012 гг., зарегистрированных в отраслевой системе.

Единственные (по сводному каталогу) экземпляры журналов, поступившие в 2011-2012 гг. распределены по фондам библиотек следующим образом:

- 89 названий хранятся в библиотеке Математического института им. В.А. Стеклова (Москва);
- 87 названий хранятся в Институте математики им. С.Л. Соболева (Новосибирск).

Две другие библиотеки в Москве С.-Петербурге имеют в 3 раза меньше поступлений (30 и 31 название соответственно) и не могут без помощи других библиотек поддерживать удовлетворительный уровень информационного обеспечения ученых. Разумеется, электронная доставка документов частично восполняет недостающие публикации в библиотеках, но не решает проблему поддержания информационной полноты фондов в крупных регионах.

Следующей частью корпоративного проекта является совместная библиографическая база данных «Статьи в отечественных математических журналах», которую ведут те же участники, кроме Института математики им. С.Л. Соболева в Новосибирске.

Начальным информационным ядром для БД оглавлений отечественных журналов был фрагмент базы данных Всероссийской книжной палаты, приобретенный Математическим институтом им. В.А. Стеклова РАН в 2003-04 гг. Этот массив включал оглавления 35 математических журналов за различные периоды времени, но не ранее 1990 г., и содержал приблизительно 40-45 тыс.

записей. В настоящее время выверенная, отредактированная и исправленная БД отражает оглавления примерно 55 основных математических журналов (с учетом изменивших название или разделенных на серии) и включает библиографические описания более чем 180 тысяч публикаций, в основном по математике, математической физике и смежным дисциплинам. Все журналы расписываются полностью вплоть до исправлений к статьям и писем в редакцию, и указывается весь авторский коллектив каждой публикации, независимо от количества авторов.

Библиографические описания журнальных публикаций вводятся в единую базу данных и отражаются в сводном журнальном каталоге 4-х библиотек, где появляется пометка, что номер журнала расписан и можно посмотреть его оглавление. Также каждая библиотека по желанию пользователя может показать на экране либо только свою часть журнального фонда, либо сводную БД, чтобы можно было найти дополнительную информацию, например, о местонахождении номеров, являющихся лакуной для данного фонда. Создание этой БД было трудоемким, но она востребована и академическими пользователями, и высшей школой. На 01.06.2013 г. зафиксировано 36562 обращения к этой базе данных, в среднем около 4,500 запросов в год.

БД создается профессиональными библиотекарями, оперативно пополняется (в течение одной недели с момента поступления номера в библиотеку), соответствует нормам библиографического описания статей, включает адекватные описания формул и находится в свободном доступе в Интернете.

Список журналов распределен между библиотеками-партнерами, и сотрудники не только осуществляют ввод новой информации в БД, но и постепенно пополняют ее более старыми выпусками журналов (углубляют ретроспективу). Руководит работой этого корпоративного консорциума заведующий библиотекой Математического института им. В.А. Стеклова, специалист по информационным системам и базам данных к.т.н. К.П. Погорелко. Он же является разработчиком программного обеспечения для всех наших проектов.

Мы полагали, что отраслевая библиографическая БД с высокой степенью надежности удовлетворит существенную часть информационных потребностей научных сотрудников и позволит в дальнейшем создать полнотекстовую БД, если редколлегия изданий и авторы публикаций захотят это сделать. Но это направление получило развитие в рамках другого корпоративного проекта, о чем будет сказано ниже.

Свободный доступ к библиографическим ресурсам осуществляется с сайта нашей библиотеки (<http://libserv.mi.ras.ru>), кнопка «Межбиблиотечный каталог», затем «Статьи в отечественных математических журналах» - и, нажав кнопку «Выбрать журнал», там можно увидеть список расписываемых журналов и последовательно через название, год, том и номер увидеть оглавление выпуска). Также возможен доступ к базе данных через наших партнеров - академический математический портал MATH-NET.RU (www.math.net.ru), доступ к которому возможен как через сайт Математического института им. В.А. Стеклова (www.mi.ras.ru), так и с сайта Президиума РАН (www.ras.ru).

Портал создавался силами Отделения математики РАН и Математического института им. В.А. Стеклова и начальной информационной базой был фонд нашей библиотеки, в частности – отечественные математические журналы, издаваемые институтом и Отделением, с первых выпусков по настоящее время. По статистике портала на 1 июня 2013 г. представлены следующие информационные ресурсы:

- 89 отечественных журналов по математике и смежным дисциплинам;
- 52624 отечественных автора (сведения о персоналиях включают краткую информацию об ученом, контактную информацию и список публикаций, отраженных в ведущих БД «Google Scholar» и «ZentralBlatt»/ Европейского математического общества);
- всего отражено 120506 публикаций, в том числе 111377 научных статей.

В связи с тем, что Министерство науки и образования придает большое значение различным рейтингам и индексам, портал предлагает возможности для сравнения, как отражается издательская активность ученого в информационном пространстве. На страницах портала можно посмотреть списки:

- часто публикуемых авторов,
- часто цитируемых авторов,
- часто цитируемых статей,
- популярных статей.

Поскольку современные издательские технологии позволяют отражать содержание новых номеров журнала еще до его выхода в свет, сроки отражения новой информации могут существенно различаться: журналы, издаваемые институтом и Отделением, появляются в БД ранее, чем, например, «Труды» и «Записки» научных учреждений из удаленных регионов. Различаются также и уровни доступа к публикациям:

- часть издателей предоставляет свободный доступ к полным текстам статей с момента выхода номера из печати;
- часть издателей запрещают смотреть полные тексты новых публикаций (как правило, этот срок составляет от одного года до 3-х лет в зависимости от соглашения);
- есть издания, полные тексты которых закрыты для внешних пользователей, но доступны из здания Математического института им. В.А. Стеклова, в том числе и из нашей библиотеки.

Информация об уровнях доступа к изданию также представлена на главной странице портала. Организаторы портала в соответствии с законом об авторских правах разрешают учителям, студентам и исследователям копирование для использования в некоммерческих целях – для образовательной и научной деятельности. Все остальные могут копировать файлы оглавлений или тексты публикаций с согласия правообладателя. Расширение круга доступных полных текстов происходит постепенно и неравномерно, т.к. в разных регионах и организациях подходы к предоставлению полных текстов в Интернете существенно различаются.

На страницах портала есть и дополнительная информация, например, о научных организациях и университетах, о научных конференциях и семинарах текущего года и представлена видеотека записей научных докладов, прозвучавших на этих мероприятиях. Эти материалы полезны и интересны не только ученым, но и преподавательскому составу математических кафедр, аспирантам и студентам, собирающимся заниматься научной работой. Есть и дополнительный сервис для ученых – зарегистрированный пользователь может создать себе личный кабинет и собирать информацию в соответствии с индивидуальными потребностями.

Если у нашей библиотеки не было возможности заключать договоры с правообладателями и получить доступ к значительной части корпуса отечественных математических сериальных изданий, то организаторы портала получили очень значительную финансовую и административную поддержку от Отделения математики РАН и хорошо справились с этой задачей, сформировав большой объем данных об отечественных сериальных изданиях.

На портале есть ссылка на нашу библиотеку (кнопка «библиотека» на главной странице), и анализируя посещения библиотечного сайта удаленными пользователями, мы видим, что около 50% посещений осуществляются пользователями портала, а вторая половина просмотров – это наши постоянные посетители.

Библиотеки РАН основным направлением в информационной деятельности считают создание универсальных по тематике или отраслевых порталов с различными возможностями поиска нужных записей, текстов, фактов. Производство электронных полнотекстовых БД долгое время не считалось задачей библиотек, поскольку за рубежом этим направлением занимаются издательства. Однако уже несколько лет Библиотека по естественным наукам (БЕН РАН) и ряд библиотек ее сети, БАН и библиотеки ее сети, Архив РАН, ИНИОН и Суперкомпьютерный центр при Президиуме РАН совместно работают над созданием электронной политематической библиотеки «Научное наследие России». Библиотеки сканируют литературу и важные исторические документы из фондов, ограничиваясь, как правило, первой третью XX века, поскольку разыскивать наследников и правообладателей, платить им за включение публикации в базу данных просто некому, да и средства РАН очень ограничены. Научная программа предполагает, что труды российских ученых XVIII-первой трети XX вв., многие из которых давно стали библиографической редкостью, в результате этой работы будут доступны всему научному сообществу, а также преподавателям и учащимся высшей школы в Интернете.

По мере расширения баз данных, все большее количество научных организаций включаются в этот культурно-исторический проект, предлагая новые объекты (целые пласты печатного и рукописного наследия) для сканирования и дальнейшего использования в электронной среде. Помимо архивов к этому проекту хотят примкнуть и некоторые музеи РАН, сохраняющие рукописное наследие крупнейших ученых России. Этот процесс «размывает» рамки первоначального проекта и создает много организационных и технических сложностей. Но он отражает и исторически близкое родство трех ветвей хранителей

культурного и научного наследия: библиотек, архивов и музеев, которые еще в XIX в. имели очень много общего.

Если проект будет продолжаться, в итоге мы получим фундаментальную общедоступную электронную библиотеку, интересную для многих университетов. Адрес корпоративной электронной библиотеки «Научное наследие России» e-heritage.ru. В настоящее время для нее оцифровано 11067 публикаций объемом 2461829 страниц. По указанному адресу пользователям доступны 9286 публикаций (1941582 стр.), остальные проходят технологическую обработку и вскоре также будут включены в электронную библиотеку и открыты для использования.

В научно-образовательных целях в библиотеке Математического института им. В.А. Стеклова РАН была создана электронная коллекция «История Российской академии наук в документах» (проект выполнялся по гранту РГНФ). Нами оцифрован ряд изданий XIX-го века, которые давно стали библиографической редкостью, например, 10-томное собрание «Материалов для истории Императорской Академии Наук», включающее документы первой половины XVIII в., двухтомная «История» П. Пекарского и 8-томная «История Российской Академии» М.И. Сухомлинова. Собрание дополнено рядом отчетов дореволюционного периода начала XX в. по физико-математическому и филологическому отделениям. Сканированные тексты отредактированы, к ним добавлены оглавления и возможность листать электронные страницы. В итоге электронная коллекция содержит интересный и разнообразный исторический материал, включающий сведения о выдающихся деятелях российской науки и культуры, который полезен и преподавателям, и учащимся, и ученым.

Коллекция была открыта для свободного доступа в Интернете на сайте нашей библиотеки (<http://libserv.mi.ras.ru>) и вызвала очень большой интерес. В первые два года было очень большое количество скачиваний всех материалов целиком, но и сейчас ею активно пользуются как профессиональные историки, так и просто любители отечественной истории. (По техническим причинам в июне 2013 г. доступ к этой электронной коллекции на сайте библиотеки закрыт и вновь откроется к осени).

Таким образом, академические библиотеки осуществляют то, что считают миссией, важнейшей общественной задачей библиотек – создают и организуют единое информационное пространство для науки и образования, доступное для всех, включающее достоверную научную информацию и сохраняющее полученное знание для будущих поколений. В настоящее время Академия наук гораздо шире открыта для сотрудничества, чем это было раньше, и мы считаем, что библиотеки вузов, и особенно университетов, имеющие базовые кафедры и аспирантуру, соответствующие тематикам научных учреждений РАН, также должны принимать участие в общих проектах. Ведь преподавательский корпус высшей школы часто включает ведущих ученых различных институтов Академии наук, и создать благоприятную информационную среду для привлечения в науку талантливой молодежи, для роста и развития отечественных научных школ – наша общая задача.